NUI Galway
OÉ Gaillimh

Spring Examinations 2010/2011

Exam Code(s)	3IF1
Exam(s)	3rd B.Sc. In Information Technology
Module Code(s)	CT332
Module(s)	Database Systems II
Paper No.	
Repeat Paper	
External Examiner(s)	Professor Michael O'Boyle
Internal Examiner(s)	Professor Gerard Lyons Dr. Jim Duggan Colm O'Riordan
Instructions	Answer any 3 questions All questions carry equal marks.
Duration	3 hours
No. of Answer books	1
No. of Pages	3
Department(s)	Information Technology

Q.1. i) Explain, with examples, the following terms:

a) Boyce Codd Normal Form
b) Non-additive join property (lossless join property)
c) De-normalisation (12)

ii) With respect to specializations in EER models, describe with a suitable example, how you might map this feature to a relational schema. (10)

iii) Explain the process of database design by synthesis. Your answer should include an explanation of functional dependencies, closure and cover sets.
Outline any disadvantages associated with this approach.
(11)

Q.2. i) Given the following fragment of a relational schema:

STUDENT:	studentno, fname, lname, address
SUBJECT:	subjectno, subjname, description, credits, duration
TAKES:	studentno, subjname, semester
LECTURER:	idno, fname, lname, office, department
DELIVERS:	subjectno, idno

Write an SQL query to return all 3 credit subjects that are delivered by a lecturer with surname (lname) "Smith" and all subjects taken by a student with surname "Smith". Represent the query as an operator tree.

Illustrate how an operator tree representing this query could be modified, using heuristic optimisation, to represent a more efficient execution strategy.

(14)

ii) Discuss the structure of a B tree and describe, with an example, the algorithm for insertion of values into a B tree. Explain, briefly, how the insertion algorithm would this differ for a B+tree.
(9)

iii) There are several approaches to hashing to a dynamic file (extendible hashing, dynamic hashing, and linear hashing). Adopting any of these approaches, show how the index structure would grow with the following key values. (Note: If adopting linear hashing, assume an initial file of one block with hash function K mod 2). You may assume two records fit in each block.

Key values: 16, 57, 32, 48, 59, 11, 77, 91, 96 (10)

Q.3. i) Given the following schedule, outline if, and how, problems might arise. Illustrate how a graph representing the conflicts can be used to show the existence of potential problems.

Ta	**Tb**	**Tc**
read_item(x)		
	read_item(y)	
	write_item(y)	
		read_item(x)
write_item(x)		
		read_item(z)
		write_item(z)
	read_item(z)	
	write_item(z)	
		write_item(x)
read_item(y)		
write_item(y)		

(7)

ii) Timestamping and two-phase locking are two approaches to ensuring conflict-serializable schedules. For either approach, show how the above schedule would continue. Show that the approach you've adopted (timestamping or two-phase locking) will guarantee conflict-serializable schedules. (17)

iii) The two-phase commit is often used in distributed databases to ensure atomicity of transactions. With respect to the different types of failure that could occur, explain the two phase commit operates to ensure atomic transactions. (9)

Q.4 i) The join operator is an important operator in relational algebra. Outline an efficient algorithm for the join operator. Illustrate its operation with a suitable example. (9)

ii) Given a parallel architecture, explain how the join operator could be implemented efficiently. Briefly, compare the efficiency of the parallel version with the algorithm suggested in part i). (9)

iii) Choosing appropriate facts, illustrate how the standard relational operators can be provided in Datalog. (9)

iv) With respect to multi-level security in databases. Define and explain the following two properties: *simple security property* and the *star property*. (6)